MEASUREMENT SENSITIVITY EQUATIONS IN ATTITUDE DETERMINATION

Mark E. Pittelkau*

Abstract

This paper focuses on constrained and unconstrained partial derivatives of the measurement equations used in spacecraft attitude determination. Six different approaches to these calculations are examined. Although all of these are found in the literature, there are many missing details and unexplained subtleties. This paper presents these details and explains the subtleties.

INTRODUCTION

Attitude determination filters are based on a model of the attitude kinematics and the measurement process. The attitude kinematics may be modeled by the quaternion differential equation and the measurement equations may be written in terms of the attitude quaternion. The quaternion must be either implicitly or explicitly constrained to unit norm for it to represent attitude. The quaternion differential equation preserves the norm of the quaternion, so for an initial quaternion $\boldsymbol{q}$ of unity norm, we have $|\boldsymbol{q}| = 1$ implicitly for all time. The measurement equation $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{q})$ is then written directly in terms of the quaternion; the constraint is *implicit* in this expression. An *explicit* unity norm constraint means that an unconstrained quaternion is explicitly normalized in the measurement equation, e.g., $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{q}/|\boldsymbol{q}|)$. Physically realizable measurements are insensitive to the norm of the quaternion, so the quaternion in the measurement model must be either implicitly or explicitly constrained.

If the quaternion norm is modeled as being arbitrary, the quaternion must be normalized either explicitly or implicitly in an attitude measurement equation. This is because any physical measurement of attitude does not depend on the quaternion norm. Nevertheless, one can find examples in the literature where physical and mathematical constraints are either ignored in the design of an attitude determination filter, or *ad hoc* procedures are included in the filter to account for constraints [1, 2]. Detailed discussions can be found in [3]–[7]. The problem of constraints is well understood in optimization and least-squares [9], but apparently not in the context of Kalman filtering and, in particular, attitude determination.

The Extended Kalman Filter requires a measurement sensitivity function, which contains the partial derivative of the measurement equation with respect to the attitude state. In this paper we focus on the quaternion as the attitude state in the filter. *This is not meant to imply approval of the quaternion as the filter state; on the contrary it is not recommended* [3]. The measurement sensitivity equation in [10, Eq. (81)] is given without mentioning subtle details of its derivation, in particular the effect of the quaternion norm constraint. It is shown in this paper that an extraneous term appears in the vector-measurement sensitivity function if the quaternion-norm constraint is disregarded when computing the partial derivative.

*The Johns Hopkins University Applied Physics Laboratory, Space Department, Laurel, MD, 20723-6099.
e-mail: `Mark.Pittelkau@jhuapl.edu`, `mpittelkau@ieee.org`

Much can be found in the literature about vector measurements, and it seems to this author that there is too much emphasis on vector measurements. The cornucopia of references on vector measurements won't be given here; the reader should find sufficient detail and references in [5]–[8]. The vector representation of an attitude measurement is generally one of convenience. In fact, only the magnetometer is a true vector sensor. Focal plane sensors, which include detector arrays and sun sensors, produce attitude measurements in two orthogonal directions in the focal plane. Devices that produce a quaternion "measurement" do so by converting several focal plane measurements to vectors, from which the quaternion is computed, usually by the QUEST algorithm [8].

We will first focus on vector measurement models and the quaternion as an unconstrained and constrained attitude state. We will then examine the focal plane measurement model, also with unconstrained and constrained quaternion attitude states. Quaternion measurement models are considered next, and finally the vector measurement model with the rotation vector as the attitude state is considered for comparison.

The notation, definitions, and identities for quaternion algebra and direction cosine matrices used in our analysis are provided in Appendix A. The equations in Appendix A are stated without proof. The studious reader can easily prove them or find them in standard texts and references such as [11] and [12], albeit with notational differences.

VECTOR MEASUREMENT, UNCONSTRAINED QUATERNION, NONNORMALIZED ATTITUDE MATRIX

The vector measurement model in this case comprises the nonnormalized attitude matrix and unconstrained quaternion,

$$\boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{v}^b = \mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i \tag{1}$$

where $\boldsymbol{v}^i$ is an inertial reference vector with $|\boldsymbol{v}^i| = 1$. This model is not physically realistic because the norm of the vector measurement should not depend on the norm of the attitude quaternion. Nevertheless, it is the model used in [1] and [2], where the norm of the quaternion is not constrained. Because of the lack of a quaternion norm constraint, the derivative is taken assuming the components of the quaternion are independent.

We will now derive the measurement sensitivity matrix. To reduce notational clutter, we will write $\boldsymbol{v}$ for $\boldsymbol{v}^i$ in equation (1). From equations (1), (A–16b), and (A–17) we have

$$\boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{v}(s^2 + |\boldsymbol{r}|^2) + 2s(\boldsymbol{v} \times \boldsymbol{r}) + 2(\boldsymbol{v} \times \boldsymbol{r}) \times \boldsymbol{r} \tag{2}$$

Define $\bar{\boldsymbol{v}} = \left[\begin{smallmatrix}\boldsymbol{v}\\0\end{smallmatrix}\right]$, $\hat{\boldsymbol{v}}^b = \boldsymbol{v}^b/|\boldsymbol{v}^b|$, $\hat{\boldsymbol{q}} = \boldsymbol{q}/|\boldsymbol{q}|$, and $\hat{\Xi} = \Xi(\hat{\boldsymbol{q}})$. (The caret denotes a unit vector, *not* an estimate.) The unconstrained measurement sensitivity matrix is obtained by differentiating equation (2) with respect to $\boldsymbol{q}$ but without regard to the norm constraint on $\boldsymbol{q}$. This sensitivity matrix is given by†

$$\begin{aligned} H_Q = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}^T} &= 2\boldsymbol{v}[\boldsymbol{r}^T \quad s] + 2\frac{\partial}{\partial \boldsymbol{q}^T}\big[s(\boldsymbol{v}\times\boldsymbol{r})\big] + 2\frac{\partial}{\partial \boldsymbol{q}^T}\big[(\boldsymbol{v}\times\boldsymbol{r})\times\boldsymbol{r}\big] \\ &= 2\boldsymbol{v}[\boldsymbol{r}^T \quad s] + 2\Big[s[\boldsymbol{v}\times] + [(\boldsymbol{v}\times\boldsymbol{r})\times] - [\boldsymbol{r}\times][\boldsymbol{v}\times] \ \Big| \ \boldsymbol{v}\times\boldsymbol{r}\Big] \\ &= 2\Big[s[\boldsymbol{v}\times] + \boldsymbol{r}\boldsymbol{v}^T - \boldsymbol{v}\boldsymbol{r}^T - [\boldsymbol{r}\times][\boldsymbol{v}\times] \ \Big| \ \boldsymbol{v}s + \boldsymbol{v}\times\boldsymbol{r}\Big] \\ &= 2[\mathbf{I} \ \ \mathbf{0}]\begin{bmatrix} s\mathbf{I} - [\boldsymbol{r}\times] & -\boldsymbol{r} \\ \boldsymbol{r}^T & s\end{bmatrix}\begin{bmatrix}[\boldsymbol{v}\times] & \boldsymbol{v} \\ -\boldsymbol{v}^T & 0\end{bmatrix} \end{aligned}$$

†The notation H_Q, H_O, and H_C is from [4].

$$= 2[\mathbf{I} \ \ \mathbf{0}]\,[\boldsymbol{q}^*\circledast]\,[\bar{\boldsymbol{v}}\circledast] \tag{3a}$$

$$= 2[\mathbf{I} \ \ \mathbf{0}]\,[\boldsymbol{q}^*\circledast]\,[\bar{\boldsymbol{v}}\circledast]\,[\boldsymbol{q}\circledast]\,[\boldsymbol{q}^*\circledast]/|\boldsymbol{q}|^2$$

$$= 2|\boldsymbol{q}|^{-2}[\mathbf{I} \ \ \mathbf{0}]\,[\bar{\boldsymbol{v}}^b\circledast]\,[\boldsymbol{q}\circledast]^T$$

$$= 2|\boldsymbol{q}|^{-2}[\mathbf{I} \ \ \mathbf{0}]\begin{bmatrix} [\boldsymbol{v}^b\times] & \boldsymbol{v}^b \\ -(\boldsymbol{v}^b)^T & 0 \end{bmatrix}\begin{bmatrix} \Xi^T \\ \boldsymbol{q}^T \end{bmatrix}$$

$$= 2|\boldsymbol{q}|^{-2}\big([\boldsymbol{v}^b\times]\Xi^T + \boldsymbol{v}^b\boldsymbol{q}^T\big) \tag{3b}$$

$$= 2\big([\hat{\boldsymbol{v}}^b\times]\hat{\Xi}^T + \hat{\boldsymbol{v}}^b\hat{\boldsymbol{q}}^T\big) \tag{3c}$$

This can be shown to be equal to Bar-Itzhack's measurement sensitivity equation [1, 2]. From equation (3a) we have

$$\begin{aligned} H_Q &= 2[\mathbf{I} \ \ \mathbf{0}]\,[\boldsymbol{q}^*\circledast]\,[\bar{\boldsymbol{v}}\circledast] \\ &= 2[\mathbf{I} \ \ \mathbf{0}]\begin{bmatrix} \Xi^T \\ \boldsymbol{q}^T \end{bmatrix}[\bar{\boldsymbol{v}}\circledast] \\ &= 2\,\Xi^T[\bar{\boldsymbol{v}}^i\circledast]\,. \end{aligned} \tag{4}$$

Multiplying this out term by term yields Bar-Itzhack's expressions for H_Q.

Shuster derived the sensitivity matrix for a vector measurement and unconstrained quaternion state [5, 6] but it is not obvious that his result is the same as equations (3a)–(3c). This equivalence is shown in Appendix B.

Because the measurement equation (1) is sensitive to the quaternion norm, the sensitivity matrix, equation (3b), contains the term $\boldsymbol{v}^b\boldsymbol{q}^T$. The sensitivity to variations of $\boldsymbol{q}$ in the direction of $\boldsymbol{q}$ is nonzero, that is, $H_Q\boldsymbol{q} = 2\boldsymbol{v}^b$, which is in the direction of the measurement vector $\boldsymbol{v}^b$.

VECTOR MEASUREMENT, UNCONSTRAINED QUATERNION, NORMALIZED ATTITUDE MATRIX

The measurement model in this case comprises the normalized attitude matrix

$$\boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{v}^b = |\boldsymbol{q}|^{-2}\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i \tag{5a}$$

$$= \mathbf{A}(\boldsymbol{q}/|\boldsymbol{q}|)\boldsymbol{v}^i\,. \tag{5b}$$

As in the previous section we will not enforce a quaternion norm constraint. The derivative is taken assuming the components of the quaternion are independent. Then

$$\begin{aligned} H_O = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}^T} &= |\boldsymbol{q}|^{-2}\frac{\partial}{\partial \boldsymbol{q}^T}\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i + \mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i\left[\frac{\partial}{\partial \boldsymbol{q}^T}|\boldsymbol{q}|^{-2}\right] \\ &= 2|\boldsymbol{q}|^{-2}\big([\boldsymbol{v}^b\times]\Xi^T + \boldsymbol{v}^b\boldsymbol{q}^T\big) - 2|\boldsymbol{q}|^{-4}\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i\boldsymbol{q}^T \\ &= 2|\boldsymbol{q}|^{-2}\big([\boldsymbol{v}^b\times]\Xi^T + \boldsymbol{v}^b\boldsymbol{q}^T\big) - 2|\boldsymbol{q}|^{-2}\boldsymbol{v}^b\boldsymbol{q}^T \\ &= 2|\boldsymbol{q}|^{-2}[\boldsymbol{v}^b\times]\Xi^T \\ &= 2[\hat{\boldsymbol{v}}^b\times]\hat{\Xi}^T\,. \end{aligned} \tag{6}$$

This does not include the term $\hat{\boldsymbol{v}}^b\hat{\boldsymbol{q}}^T$ seen in equation (3c) because the quaternion is explicitly normalized in the measurement equation (5), so there is no sensitivity to variations of $\boldsymbol{q}$ in the direction of $\boldsymbol{q}$, that is, $H_O\boldsymbol{q} = \mathbf{0}$.

VECTOR MEASUREMENT, CONSTRAINED QUATERNION

We now examine the case where the quaternion is fixed to unit norm in the vector measurement equation. Because the quaternion norm is now fixed, there is no variation in the direction of $\boldsymbol{q}$. The measurement sensitivity matrix can be derived by projecting H_Q onto the orthogonal complement of $\hat{\boldsymbol{q}}$. Postmultiplying (3c) by $\boldsymbol{I} - \hat{\boldsymbol{q}}\hat{\boldsymbol{q}}^T = \hat{\Xi}\,\hat{\Xi}^T$ yields

$$\begin{aligned} H_C &= H_Q\,\hat{\Xi}\,\hat{\Xi}^T \\ &= 2[\hat{\boldsymbol{v}}^b\times]\hat{\Xi}^T \,, \end{aligned} \tag{7}$$

which is the same as (6). This is the "classical" (LMS) definition [10]. Note that $H_C\boldsymbol{q} = \boldsymbol{0}$, which should already be obvious.

FOCAL PLANE MEASUREMENTS

The vector $\boldsymbol{v}^s$ in the sensor reference frame is related to a vector $\boldsymbol{v}^i$ in the inertial reference frame i by

$$\boldsymbol{v}^s = \mathbf{T}_b^s\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i \,, \tag{8}$$

where $\mathbf{T}_b^s$ is the transformation (mounting) matrix from the body frame to the sensor frame. ($\mathbf{T}_b^s$ was omitted from equation (1), without loss of generality, to simplify the discussion.) We can also consider the explicitly normalized model

$$\boldsymbol{v}^s = |\boldsymbol{q}|^{-2}\mathbf{T}_b^s\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i \,. \tag{9}$$

The line-of-sight vector in the sensor frame is $\boldsymbol{v}^s = [v_x^s,\, v_y^s,\, v_z^s]^T$. The focal-plane measurement model is

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{v}^s) = \begin{bmatrix} v_x^s/v_z^s \\ v_y^s/v_z^s \end{bmatrix} \,. \tag{10}$$

Our analysis will start with equations (8) and (10) for the case of an unconstrained quaternion state. The measurement sensitivity matrix is obtained by the chain rule

$$H_F = \frac{\partial \boldsymbol{h}(\boldsymbol{v}^s)}{\partial \boldsymbol{q}^T} = \frac{\partial \boldsymbol{h}}{\partial (\boldsymbol{v}^s)^T}\frac{\partial \boldsymbol{v}^s}{\partial \boldsymbol{q}^T} \,. \tag{11}$$

Differentiating (10) and substituting (3c) we get

$$\begin{aligned} H_F &= \begin{bmatrix} 1/v_z^s & 0 & -v_x^s/(v_z^s)^2 \\ 0 & 1/v_z^s & -v_x^s/(v_z^s)^2 \end{bmatrix} \mathbf{T}_b^s\left[2\big([\hat{\boldsymbol{v}}^b\times]\hat{\Xi}^T + \hat{\boldsymbol{v}}^b\hat{\boldsymbol{q}}^T\big)\right] \\ &= \begin{bmatrix} 1/v_z^s & 0 & -v_x^s/(v_z^s)^2 \\ 0 & 1/v_z^s & -v_x^s/(v_z^s)^2 \end{bmatrix} \left[2\big([\hat{\boldsymbol{v}}^s\times]\,\mathbf{T}_b^s\,\hat{\Xi}^T + \hat{\boldsymbol{v}}^s\hat{\boldsymbol{q}}^T\big)\right] \\ &= \begin{bmatrix} 1/v_z^s & 0 & -v_x^s/(v_z^s)^2 \\ 0 & 1/v_z^s & -v_x^s/(v_z^s)^2 \end{bmatrix} [\hat{\boldsymbol{v}}^s\times]\,\mathbf{T}_b^s\,\hat{\Xi}^T \end{aligned} \tag{12}$$

The middle step was obtained by using equation (A–4) and the term $\hat{\boldsymbol{v}}^s\hat{\boldsymbol{q}}^T$ was annihilated in the last step because variations in the direction of $\hat{\boldsymbol{v}}^s$ (caused by norm variations of $\boldsymbol{q}$) do not affect the focal plane measurement as can be seen in equation (10). Equation (12) is obtained also if we start with the vector measurement model (8) for the constrained quaternion state or (9) for the unconstrained quaternion state. Thus we see that the sensitivity to variations in the quaternion norm is zero, i.e., $H_F\boldsymbol{q} = \boldsymbol{0}$.

QUATERNION MEASUREMENTS

There is no device that physically measures a quaternion, rather it is computed from vector measurements or from focal plane measurements converted to vectors, usually by using the QUEST algorithm [8]. The computed quaternion "measurement" $\boldsymbol{p}$ has unit norm, so the measurement as a function of an unconstrained quaternion state $\boldsymbol{q}$ is

$$\boldsymbol{p} = \boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{q}/|\boldsymbol{q}| \,. \tag{13}$$

The measurement sensitivity matrix is then

$$\begin{aligned} H_P &= \frac{\partial}{\partial \boldsymbol{q}^T}\big(\boldsymbol{q}(\boldsymbol{q}^T\boldsymbol{q})^{-1/2}\big) \\ &= \frac{1}{|\boldsymbol{q}|}\left(\boldsymbol{I} - \frac{1}{|\boldsymbol{q}|^2}\boldsymbol{q}\boldsymbol{q}^T\right) \\ &= \frac{1}{|\boldsymbol{q}|}\,\hat{\Xi}\,\hat{\Xi}^T \end{aligned} \tag{14}$$

Since the quaternion "measurement" $\boldsymbol{p}$ is of unit norm, the sensitivity of $\boldsymbol{p}$ to variations of the state $\boldsymbol{q}$ in the direction of $\boldsymbol{q}$ is zero, that is, $H_P\boldsymbol{q} = \boldsymbol{0}$. For the measurement model $\boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{q}$ with the norm of $\boldsymbol{q}$ *constrained*, the sensitivity matrix is given by equation (14) and we get $H_P\boldsymbol{q} = \boldsymbol{0}$.

In [1] and [2] the quaternion measurement equation is taken as $\boldsymbol{h}(\boldsymbol{q}) = \boldsymbol{q}$, where the quaternion norm is unconstrained, and the sensitivity matrix there is $H = \boldsymbol{I}$ (the 4×4 identity matrix), which is incorrect.

MEASUREMENT EQUATION OF CHOUKROUN, ET AL.

Choukroun, et al., [13] developed the derived measurement equation

$$\boldsymbol{z} = \bar{\boldsymbol{w}}^b \otimes \boldsymbol{q} - \boldsymbol{q} \otimes \bar{\boldsymbol{v}}^i \,, \tag{15}$$

where $\boldsymbol{v}^i$ is an inertial reference vector, $\boldsymbol{w}^b$ is a vector measurement in the body frame, and $\boldsymbol{z}$ is the derived measurement. (For convenience, the sign of $\boldsymbol{z}$ is changed from that in [13], which is inconsequential because the expression equals zero.) This measurement equation originated with Shuster [5] in the context of this paper, and was derived independently for a square-root QUEST algorithm in [14].

The equivalence of equations (15) and (5) can be seen in the following result.

$$\begin{aligned} \boldsymbol{z} &= \bar{\boldsymbol{w}}^b \otimes \boldsymbol{q} - \boldsymbol{q} \otimes \bar{\boldsymbol{v}}^i \\ &= [\boldsymbol{q}\circledast]\bar{\boldsymbol{w}}^b - [\boldsymbol{q}\otimes]\bar{\boldsymbol{v}}^i \\ &= [\boldsymbol{q}\circledast]\big(\bar{\boldsymbol{w}}^b - |\boldsymbol{q}|^{-2}[\boldsymbol{q}\circledast]^T[\boldsymbol{q}\otimes]\bar{\boldsymbol{v}}^i\big) \\ &= \Xi\big(\boldsymbol{w}^b - |\boldsymbol{q}|^{-2}\mathbf{A}(\boldsymbol{q})\boldsymbol{v}^i\big) \,. \end{aligned} \tag{16}$$

The expression in parenthesis is the filter residual; the leading Ξ disappears upon substitution into the Kalman filter (attitude determination filter) equations, leaving a multiplier of $|\boldsymbol{q}|^2$ in the gain equation (because $\Xi^T\Xi = |\boldsymbol{q}|^2\mathbf{I}$), which can be significant if $|\boldsymbol{q}|^2$ (the norm-squared of the estimated quaternion) is not close to unity.

ROTATION VECTOR ATTITUDE STATE

Attitude is inherently defined as a rotation with a minimum of three parameters. All three-dimensional parameterizations of rotation exhibit a singularity for some particular angle of rotation, and it is for this reason that the four-parameter quaternion is popular, even though it is subject to a unit-norm constraint. Theoretical and practical complications of the quaternion attitude estimator, which are due to the norm constraint and the 2-1 redundancy of the quaternion attitude representation, are the subject of [3]–[7] and are not addressed here. The singularity in the three-parameter representation can be avoided by ensuring that the angle of rotation is always small [15].

We have thus far examined the measurement sensitivity equations when the attitude state is a quaternion. For comparison, we will now look at the measurement sensitivity equation when the attitude is parameterized by a rotation vector [15]. The vector measurement equation can be written

$$\boldsymbol{v}^s = \mathbf{T}_b^s \mathbf{A}(\boldsymbol{\phi}_i^b)\boldsymbol{v}^i \,, \tag{17}$$

where the attitude matrix is parameterized by the rotation vector as shown in equation (A–18a). There is no constraint, and the measurement sensitivity matrix is easily found to be

$$H_{\boldsymbol{\phi}} = \frac{\partial \boldsymbol{v}^s}{\partial \boldsymbol{\phi}^T} = \mathbf{T}_b^s[\boldsymbol{v}^b \times] \,. \tag{18}$$

CONCLUSION

This paper presented many details and explained many subtleties that arise in computing measurement sensitivity matrices for the purpose of attitude determination when the quaternion is the attitude state. Many of these details and subtleties are not found or explained in the literature.

The quaternion must be normalized to unit norm either explicitly or implicitly in the measurement equation for the measurement equation to be a physically and mathematically valid model. Failure to do so can result in extraneous terms in the measurement sensitivity function, depending on whether the measurement is a vector, focal plane coordinates, or a normalized quaternion. Furthermore, failure to recognize the quaternion constraint in the measurement model, regardless of whether it is constrained in the filter state, can lead to an incorrect measurement sensitivity equation. Inconsistent or ambiguous (i.e., incorrect) filter state estimates and covariances can then result [3, 4], thus leading to *ad hoc* fixes and arbitrary "tuning" to make the quaternion attitude determination filter work.

In the measurement model (1), the quaternion must be constrained to unit norm to obtain a valid vector measurement. In the measurement model (5), the quaternion kinematics can be unconstrained, but normalization appears explicitly in the measurement model to obtain a valid vector measurement. Thus, (6) is the correct measurement sensitivity matrix for vector measurements; equations (3a)–(3c) are not the correct expressions to use.

Focal plane measurements are insensitive to the quaternion norm, and it was shown that the extraneous term in equation (3c) is annihilated in the sensitivity matrix for focal plane measurements. If the quaternion constraint is not properly handled, the filter state estimate can behave differently for different types of attitude measurements containing the same attitude information.

The measurement sensitivity matrix for a quaternion measurement is commonly taken to be the identity matrix, but because the measurement is of unit norm, the quaternion measurement sensitivity matrix is not the identity matrix, regardless of whether the filter quaternion state is constrained.

ACKNOWLEDGEMENT

This article was motivated by discussions with Malcolm Shuster and F. Landis Markley about constrained quaternion estimation. Malcolm Shuster has written an in-depth two-part treatment of constrained and unconstrained quaternion estimation [6, 7] that encompasses some of the results of this article and is an extension of earlier work [5]. Landis Markley has produced analytical results comparing unconstrained quaternion estimation and attitude estimation parameterized by three small rotation angles [4]. His analysis also encompasses some of the results of this article.

REFERENCES

1. BAR-ITZHACK, I. Y., and OSHMAN, Y., "Attitude Determination from Vector Measurements: Quaternion Estimation", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-21, No. 1, 1985, pp. 128–136.
2. BAR-ITZHACK, I. Y. and THIENEL, J., "On the Singularity in the Estimation of the Quaternion-of-Rotation", Paper Number AIAA-2002-4831, AIAA Astrodynamics Specialist Conference, Monterey, CA, 8 August 2002, pp. 931–941.
3. PITTELKAU, M. E., "An Analysis of the Quaternion Attitude Determination Filter", AAS/AIAA Space Flight Mechanics Meeting, Paper No. AAS 03-194, in *Advances of the Astronautical Sciences*, Vol. 114, Part II, pp. 1337–1352, 2003. Accepted for publication in *Journal of the Astronautical Sciences*.
4. MARKLEY, F. L., "Attitude Estimation or Quaternion Estimation?", AAS J. L. Junkins Astrodynamics Symposium, 2003, and in the NASA/GSFC Flight Mechanics Symposium, 28–30 October 2003, (also to appear in a special issue of the JAS).
5. SHUSTER, M. D., "The Quaternion in the Kalman Filter", *Advances in the Astronautical Sciences*, Paper No. AAS-93-553, Vol. 85, 1993, pp. 25–37.
6. SHUSTER, M. D., "Constraint In Attitude Estimation, Part I: Constrained Estimation", Submitted to the Journal of the Astronautical Sciences, 2003.
7. SHUSTER, M. D., "Constraint In Attitude Estimation, Part II: Unconstrained Estimation", Submitted to the Journal of the Astronautical Sciences, 2003.
8. SHUSTER, M. D. and OH, S. D., "Three-Axis Attitude Determination from Vector Observations", *AIAA Journal of Guidance and Control*, Vol. 4, No. 1, Jan-Feb 1981, pp. 70–77.
9. GILL, P. E., MURRAY, W., WRIGHT, M. H., *Practical Optimization*, Academic Press, New York, 1981.
10. LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D., "Kalman Filtering for Spacecraft Attitude Estimation", *AIAA J. Guidance, Control, Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429.
11. WERTZ, J. R., ed., *Spacecraft Attitude Determination and Control*, Kluwer Academic Publishers, 1978, pp. 410–416, 510–512, 564–565, 755–762.
12. SHUSTER, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439-517.
13. CHOUKROUN, D., BAR-ITZHACK, I. Y., and OSHMAN, Y., "A Novel Quaternion Kalman Filter", Paper AIAA-2002-4460, AIAA Guidance, Navigation, and Control Conference, Monteray, CA, 8 August 2002, Vol. 1, pp. 219–229.
14. PITTELKAU, M. E., "Square Root Quaternion Estimation", Paper AIAA-2002-4914, *AIAA Guidance, Navigation, and Control Conference*, Monteray, CA, 8 August 2002.
15. PITTELKAU, M. E., "The Rotation Vector In Attitude Determination", *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 6, 2003. (accepted for publication).

A NOTATION, DEFINITIONS, AND QUATERNION PROPERTIES

Notation, definitions, quaternion properties and identities, and direction cosine matrix (DCM) identities are provided in this appendix. We will denote a 3×3 identity matrix by $\mathbf{I}$ and all other identity matrices by $\boldsymbol{I}$. A vector $\boldsymbol{v}$ in a particular frame a is denoted by $\boldsymbol{v}^a$. A transformation matrix is a DCM that can be used to transform a vector from a frame a to a frame b, which we will denote by $\mathbf{T}_a^b$, so we have $\boldsymbol{v}^b = \mathbf{T}_a^b \boldsymbol{v}^a$. Quaternions may be adorned in the same manner as transformation matrices.

Cross Product Operator and Operator Matrix

he cross-product operator $\times$ is such that for vectors $\boldsymbol{u} = [u_x,\, u_y,\, u_z]^T$ and $\boldsymbol{v} = [v_x,\, v_y,\, v_z]^T$, we have

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{bmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{bmatrix} . \tag{A–1}$$

The cross-product operator matrix $[\boldsymbol{u}\times]$ is defined such that

$$[\boldsymbol{u}\times]\boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{v} \tag{A–2}$$

where

$$[\boldsymbol{u}\times] = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} . \tag{A–3}$$

Since the cross-product matrix is skew-symmetric, we have $[\boldsymbol{u}\times]^T = -[\boldsymbol{u}\times]$ and $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$. A useful identity is $[\boldsymbol{u}\times]^2 = |\boldsymbol{u}|^2 \mathbf{I} - \boldsymbol{u}\boldsymbol{u}^T$. A reference frame transformation may be applied to the cross-product matrix so that

$$[\mathbf{T}_a^b \boldsymbol{v}^a \times] = \mathbf{T}_a^b [\boldsymbol{v}^a \times] \mathbf{T}_b^a . \tag{A–4}$$

Quaternion Operators

In the definitions that follow, let $\boldsymbol{p}$, $\boldsymbol{q}$, and $\boldsymbol{r}$ be quaternions represented as 4×1 column matrices and let $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{q}}$, and $\tilde{\boldsymbol{r}}$ be the corresponding hypercomplex quaternions. Also let $\mathbf{T}_{\boldsymbol{p}}$, $\mathbf{T}_{\boldsymbol{q}}$, and $\mathbf{T}_{\boldsymbol{r}}$ be the corresponding orthonormal direction cosine matrices. The subscript here is a shorthand notation that simply indicates that $\mathbf{T}$ corresponds to a particular quaternion, otherwise we would have to write, for example, $\mathbf{T}(\boldsymbol{q})$.

Quaternion operator notation is necessary because while the hypercomplex product $\tilde{\boldsymbol{p}}\tilde{\boldsymbol{q}}$ is well-defined, it is technically incorrect and sometimes causes confusion to write the matrix product $\boldsymbol{p}\boldsymbol{q}$. The use of operator notation removes this difficulty and provides much greater flexibility in performing quaternion algebra. The quaternion operators and operator matrices to be introduced are analagous to the cross-product operator $\times$ and operator matrix $[\boldsymbol{u}\times]$. The quaternion $\boldsymbol{q} = [q_x,\, q_y,\, q_z,\, q_s]^T$ comprises the vector $[q_x,\, q_y,\, q_z]^T$ and the scalar q_s. The norm (2-norm) of a quaternion is given by $|\boldsymbol{q}| = (q_x^2 + q_y^2 + q_z^2 + q_s)^{1/2}$. It is usually assumed in attitude work that quaternions have unit norm, but for generality we will not make that assumption here. (Ultimately, however, an implicit or explicit normalization occurs when computing a rotation vector, Gibbs or Rodriguez vector, DCM, or any other attitude representation.) The conjugate $\boldsymbol{q}^*$ of a quaternion $\boldsymbol{q} = [q_x,\, q_y,\, q_z,\, q_s]^T$ is $\boldsymbol{q}^* = [-q_x,\, -q_y,\, -q_z,\, q_s]^T$. The inverse of $\boldsymbol{q}$ is $\boldsymbol{q}^{-1} = \boldsymbol{q}^*/|\boldsymbol{q}|^2$ so that $\boldsymbol{q} \otimes \boldsymbol{q}^{-1} = \boldsymbol{q}^{-1} \otimes \boldsymbol{q} = [0, 0, 0, 1]^T$. For a unit quaternion, $\boldsymbol{q}^* = \boldsymbol{q}^{-1}$ so that $\boldsymbol{q}^* \otimes \boldsymbol{q} = \boldsymbol{q} \otimes \boldsymbol{q}^* = [0, 0, 0, 1]^T$ and similarly for the $\circledast$ operator, where $\otimes$ and $\circledast$ are quaternion multiplication operators defined below. The $[\boldsymbol{q}\otimes]$ and $[\boldsymbol{q}\circledast]$ operator matrices are also introduced. The $\otimes$ operator was introduced in [10]. The operator matrices are this author's invention.

The $\otimes$ Operator

The operator $\otimes$ and operator matrix $[\boldsymbol{q}\otimes]$ are defined by

$$\begin{aligned}\boldsymbol{r} &= \boldsymbol{p}\otimes\boldsymbol{q} \\ &= [\boldsymbol{p}\otimes]\boldsymbol{q} \\ &= \begin{bmatrix} p_s & p_z & -p_y & p_x \\ -p_z & p_s & p_x & p_y \\ p_y & -p_x & p_s & p_z \\ -p_x & -p_y & -p_z & p_s \end{bmatrix}\begin{bmatrix} q_x \\ q_y \\ q_z \\ q_s \end{bmatrix}\end{aligned} \tag{A–5}$$

For unit-norm quaternions, the quaternion product corresponds to the DCM product

$$\boldsymbol{r} = \boldsymbol{p}\otimes\boldsymbol{q} \Longleftrightarrow \mathbf{T}_{\boldsymbol{r}} = \mathbf{T}_{\boldsymbol{p}}\mathbf{T}_{\boldsymbol{q}}$$

where the order of corresponding terms is the same, whereas in comparison with the hypercomplex product

$$\boldsymbol{p}\otimes\boldsymbol{q} \Longleftrightarrow \tilde{\boldsymbol{q}}\tilde{\boldsymbol{p}}$$

the order of corresponding terms is reversed. Define

$$\varphi = |\boldsymbol{\phi}| \qquad \bar{\boldsymbol{\phi}} = \begin{bmatrix} \boldsymbol{\phi} \\ 0 \end{bmatrix} \qquad \boldsymbol{q} = \begin{bmatrix} \frac{1}{2}\boldsymbol{\phi}\,\dfrac{\sin(\varphi/2)}{\varphi/2} \\ \cos(\varphi/2) \end{bmatrix} \tag{A–6}$$

Some useful identities follow:

$$(\boldsymbol{p}\otimes\boldsymbol{q})^* = \boldsymbol{q}^*\otimes\boldsymbol{p}^* \tag{A–7a}$$

$$[\boldsymbol{q}\otimes][\boldsymbol{q}^*\otimes] = [\boldsymbol{q}\otimes][\boldsymbol{q}\otimes]^T = |\boldsymbol{q}|^2\boldsymbol{I} \tag{A–7b}$$

$$[\boldsymbol{q}^{-1}\otimes] = [\boldsymbol{q}\otimes]^{-1} = [\boldsymbol{q}\otimes]^T/|\boldsymbol{q}|^2 \tag{A–7c}$$

$$[\boldsymbol{q}\otimes] = \exp\{\tfrac{1}{2}[\bar{\boldsymbol{\phi}}\otimes]\} = \cos(\varphi/2)\boldsymbol{I} + \tfrac{1}{2}[\bar{\boldsymbol{\phi}}\otimes]\,\frac{\sin(\varphi/2)}{\varphi/2} \tag{A–7d}$$

The $\circledast$ Operator

The $\circledast$ operator and operator matrix $[\boldsymbol{q}\circledast]$ are defined by

$$\begin{aligned}\boldsymbol{r} &= \boldsymbol{q}\circledast\boldsymbol{p} \\ &= [\boldsymbol{q}\circledast]\boldsymbol{p} \\ &= \begin{bmatrix} q_s & -q_z & q_y & q_x \\ q_z & q_s & -q_x & q_y \\ -q_y & q_x & q_s & q_z \\ -q_x & -q_y & -q_z & q_s \end{bmatrix}\begin{bmatrix} p_x \\ p_y \\ p_z \\ p_s \end{bmatrix}\end{aligned} \tag{A–8}$$

For unit-norm quaternions, the quaternion product corresponds to the DCM product

$$\boldsymbol{r} = \boldsymbol{q}\circledast\boldsymbol{p} \Longleftrightarrow \mathbf{T}_{\boldsymbol{r}} = \mathbf{T}_{\boldsymbol{p}}\mathbf{T}_{\boldsymbol{q}} \tag{A–9}$$

where the order of corresponding terms is reversed, whereas in comparison with the hypercomplex product

$$\boldsymbol{q} \circledast \boldsymbol{p} \Longleftrightarrow \tilde{\boldsymbol{q}}\tilde{\boldsymbol{p}} \tag{A–10}$$

the order of corresponding terms is the same. Some useful identities are

$$(\boldsymbol{q} \circledast \boldsymbol{p})^* = \boldsymbol{p}^* \circledast \boldsymbol{q}^* \tag{A–11a}$$

$$[\boldsymbol{q}\circledast][\boldsymbol{q}^*\circledast] = [\boldsymbol{q}\circledast][\boldsymbol{q}\circledast]^T = |\boldsymbol{q}|^2\boldsymbol{I} \tag{A–11b}$$

$$[\boldsymbol{q}^{-1}\circledast] = [\boldsymbol{q}\circledast]^{-1} = [\boldsymbol{q}\circledast]^T/|\boldsymbol{q}|^2 \tag{A–11c}$$

$$[\boldsymbol{q}\circledast] = \exp\{\tfrac{1}{2}[\bar{\boldsymbol{\phi}}\circledast]\} = \cos(\varphi/2)\boldsymbol{I} + \tfrac{1}{2}[\bar{\boldsymbol{\phi}}\circledast]\frac{\sin(\varphi/2)}{\varphi/2} \tag{A–11d}$$

$$[\boldsymbol{q}\otimes][\boldsymbol{q}\circledast]^{-1} = [\boldsymbol{q}\circledast]^{-1}[\boldsymbol{q}\otimes] = \begin{bmatrix} \mathbf{T}_{\boldsymbol{q}} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{A–11e}$$

$$\boldsymbol{p} \otimes \boldsymbol{q} = \boldsymbol{q} \circledast \boldsymbol{p} \tag{A–11f}$$

where φ, $\bar{\boldsymbol{\phi}}$, and $\boldsymbol{q}$ as a function of $\boldsymbol{\phi}$ were defined in equation (A–6).

The $[\boldsymbol{q}\circledast]$ matrix can be partitioned such that $[\boldsymbol{q}\circledast] = [\Xi \;\; \boldsymbol{q}]$, where Ξ is a 4×3 matrix. We will write $\Xi(\boldsymbol{p})$ for the partition of $[\boldsymbol{p}\circledast]$, but otherwise assume that Ξ without an argument depends on $\boldsymbol{q}$ in order to simplify the notation. Since $[\boldsymbol{q}\circledast]$ is an orthonormal matrix for $|\boldsymbol{q}| = 1$, we have that $\Xi\Xi^T + \boldsymbol{q}\boldsymbol{q}^T = |\boldsymbol{q}|^2\boldsymbol{I}$, $\Xi^T\Xi = |\boldsymbol{q}|^2\mathbf{I}$, and $\Xi^T\boldsymbol{q} = \mathbf{0}$.

Quaternion Transformation of a Vector

Let $\boldsymbol{v}^a$ be a vector in a frame a and $\boldsymbol{v}^b$ be a vector in a frame b. If a quaternion $\boldsymbol{q}$ represents the attitude of frame b with respect to frame a, the $\boldsymbol{v}^a$ is transformed to the frame b by

$$\bar{\boldsymbol{v}}^b = \boldsymbol{q} \otimes \bar{\boldsymbol{v}}^a \otimes \boldsymbol{q}^{-1} \tag{A–12a}$$

$$= \boldsymbol{q}^{-1} \circledast \bar{\boldsymbol{v}}^a \circledast \boldsymbol{q} \tag{A–12b}$$

where $\bar{\boldsymbol{v}}^a = \begin{bmatrix} \boldsymbol{v}^a \\ 0 \end{bmatrix}$ and $\bar{\boldsymbol{v}}^b = \begin{bmatrix} \boldsymbol{v}^b \\ 0 \end{bmatrix}$. The equivalence of equations (A–12a) and (A–12b) follows from equation (A–11f). From equation (A–11e), we have that equations (A–12a) and (A–12b) are the same as

$$\begin{bmatrix} \boldsymbol{v}^b \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{\boldsymbol{q}} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^a \\ 0 \end{bmatrix}. \tag{A–13}$$

Associativity and Commutativity

Quaternion operators of the same kind are associative, that is,

$$\boldsymbol{p} \otimes \boldsymbol{q} \otimes \boldsymbol{r} = (\boldsymbol{p} \otimes \boldsymbol{q}) \otimes \boldsymbol{r} = \boldsymbol{p} \otimes (\boldsymbol{q} \otimes \boldsymbol{r}) \tag{A–14a}$$

$$\boldsymbol{p} \circledast \boldsymbol{q} \circledast \boldsymbol{r} = (\boldsymbol{p} \circledast \boldsymbol{q}) \circledast \boldsymbol{r} = \boldsymbol{p} \circledast (\boldsymbol{q} \circledast \boldsymbol{r}) \tag{A–14b}$$

However, quaternion operators that are not the same kind are not associative, that is,

$$(\boldsymbol{p} \otimes \boldsymbol{q}) \circledast \boldsymbol{r} \neq \boldsymbol{p} \otimes (\boldsymbol{q} \circledast \boldsymbol{r}) \tag{A–15a}$$

$$(\boldsymbol{p} \circledast \boldsymbol{q}) \otimes \boldsymbol{r} \neq \boldsymbol{p} \circledast (\boldsymbol{q} \otimes \boldsymbol{r}) \tag{A–15b}$$

so the expressions $\boldsymbol{p} \otimes \boldsymbol{q} \circledast \boldsymbol{r}$ and $\boldsymbol{p} \circledast \boldsymbol{q} \otimes \boldsymbol{r}$ are ambiguous.

Equation (A–11f) indicates that quaternions are, in general, not commutative in their product. They are commutative only if their axes of rotation are in the same direction, which is also true of direction cosine matrices.

Parameterizations of the Attitude Matrix

Equation (A–11e) shows one way to compute the attitude matrix $\mathbf{A}(\boldsymbol{q}) = \mathbf{T}_{\boldsymbol{q}}$, namely

$$\begin{aligned} \mathbf{A}(\boldsymbol{q}) &= (s^2 - |\boldsymbol{r}|^2)\mathbf{I} - 2s[\boldsymbol{r}\times] + 2\boldsymbol{r}\boldsymbol{r}^T && \text{(A–16a)} \\ &= (s^2 + |\boldsymbol{r}|^2)\mathbf{I} - 2s[\boldsymbol{r}\times] + 2[\boldsymbol{r}\times]^2 && \text{(A–16b)} \end{aligned}$$

where the quaternion is partitioned into its vector part $\boldsymbol{r}$ and a scalar part s,

$$\boldsymbol{q} = \begin{bmatrix} \boldsymbol{r} \\ s \end{bmatrix} . \tag{A–17}$$

The attitude matrix can also be written in terms of a rotation vector $\boldsymbol{\phi}$ with $\varphi = |\boldsymbol{\phi}|$,

$$\begin{aligned} \mathbf{A}(\boldsymbol{\phi}) = \exp(-[\boldsymbol{\phi}\times]) &= (\cos\varphi)\mathbf{I} - \frac{\sin\varphi}{\varphi}[\boldsymbol{\phi}\times] + \frac{1-\cos\varphi}{\varphi^2}\boldsymbol{\phi}\boldsymbol{\phi}^T && \text{(A–18a)} \\ &= \mathbf{I} - \frac{\sin\varphi}{\varphi}[\boldsymbol{\phi}\times] + \frac{1-\cos\varphi}{\varphi^2}[\boldsymbol{\phi}\times]^2 && \text{(A–18b)} \end{aligned}$$

B MEASUREMENT SENSITIVITY MATRIX (SHUSTER)

A different approach is taken in [5, 6] to compute the sensitivity of a vector measurement to the unconstrained quaternion state. However, the result obtained therein (Eq. (37) of [5] is of a different form than equations (3a) and (3c) and it is not obvious that it is equivalent to equations (3a) and (3c). We will first outline the derivation from [5, 6] (using this author's notation) and then show how to produce the same result as (3a). We start with Eq. (17) of [5], which is the scalar measurement

$$z = \boldsymbol{u}^T \mathbf{A}(\boldsymbol{q}) \boldsymbol{v}^i \tag{B–1}$$

where $\boldsymbol{v}^i$ is a reference vector, $\mathbf{A}(\boldsymbol{q})$ is the attitude matrix computed from the attitude quaternion $\boldsymbol{q}$, and $\boldsymbol{u}$ is an auxilary vector. Substitute equation (A–16b) into the measurement equation and differentiate with respect to $\boldsymbol{q}$. The algebra involved is similar to that in the first section after the Introduction to this paper. The measurement sensitivity matrix is

$$H_{\boldsymbol{q}} = \frac{\partial z}{\partial \boldsymbol{q}^T} = -2(\bar{\boldsymbol{u}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}})^T \tag{B–2}$$

By alternately letting $\boldsymbol{u}$ equal the elementary vectors $\mathbf{1} = [1, 0, 0]^T$, $\mathbf{2} = [0, 1, 0]^T$, and $\mathbf{3} = [0, 0, 1]^T$, we construct the complete vector measurement equation (1) and the measurement sensitivity matrix

$$H_{\boldsymbol{q}} = -2 \begin{bmatrix} (\bar{\mathbf{1}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}})^T \\ (\bar{\mathbf{2}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}})^T \\ (\bar{\mathbf{3}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}})^T \end{bmatrix} \tag{B–3}$$

from which we obtain almost immediately,

$$H_{\boldsymbol{q}} = -2\, \Xi^T (\boldsymbol{q} \otimes \bar{\boldsymbol{v}}) . \tag{B–4}$$

In this expression, $\boldsymbol{q} \otimes \bar{\boldsymbol{v}}$ is the argument of Ξ, that is, Ξ is formed from the elements of $\boldsymbol{q} \otimes \bar{\boldsymbol{v}}$. Equation (B–4) does not appear to be the same as equation (3a), but its equivalence is easily established as follows. Transpose equation (B–3) and apply equation (A–11f) to get

$$\begin{aligned} H_{\boldsymbol{q}}^T &= -2\Big[\bar{\mathbf{1}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}} \;\Big|\; \bar{\mathbf{2}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}} \;\Big|\; \bar{\mathbf{3}} \otimes \boldsymbol{q} \otimes \bar{\boldsymbol{v}}\Big] \\ &= -2\Big[\bar{\boldsymbol{v}} \circledast \boldsymbol{q} \circledast \bar{\mathbf{1}} \;\Big|\; \bar{\boldsymbol{v}} \circledast \boldsymbol{q} \circledast \bar{\mathbf{2}} \;\Big|\; \bar{\boldsymbol{v}} \circledast \boldsymbol{q} \circledast \bar{\mathbf{3}}\Big] \\ &= -2[\bar{\boldsymbol{v}}\circledast]\,[\boldsymbol{q}\circledast]\begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \end{aligned} \tag{B–5}$$

Transposing again yields

$$H_{\boldsymbol{q}} = 2[\mathbf{I} \;\; \mathbf{0}]\,[\boldsymbol{q}^*\circledast]\,[\bar{\boldsymbol{v}}\circledast] \tag{B–6}$$

which is the same as equation (3a).